CHAPTER I

SCALAR QUANTITIES AND VECTOR QUANTITIES

1·1 Scalar quantities and the uniform scale. In Natural Philosophy we study classes of physical quantities that are capable of measurement—classes of quantities such as masses, lengths, velocities. The simplest classes are those for which the members are specified, on an appropriate scale, by a single number—for example, masses, lengths, temperatures. We begin with a sketch of the theory of the measurement.

We consider, then, a class K of such quantities. Our first object is to set up a scale of measurement. The possibility of such a scale depends on the *axiom of congruence*, i.e. the assumption that there is a meaning to the statement that two particular members of K are equal, or congruent, even before any scale has been set up. For example, we assume that there is a meaning to the statement that the mass of a body A is equal to the mass of a body B, and this equality is absolute, something existing prior to the construction of a scale.

The axiom of congruence is all that we need in order to set up a scale. Each member of K is labelled with a measure number, or *measure*, and in the first instance the only restriction on the measures is that congruent members of K have the same measure. Two congruent members of K are considered to be identical and interchangeable physical quantities (so far as the quantity under discussion is concerned) and the measure is taken to be a complete and adequate description. For example, any two particles of equal mass are considered to be identical for purposes of dynamics.

But any useful scale must possess two other properties, which we may call *continuity* and *direction*. We assume that we can recognize when two members of K are nearly equal, and we naturally choose a scale in which nearly equal members have nearly equal measures. Such a scale has the property of continuity. We assume that we can recognize which of two members of K has the greater magnitude, and we naturally choose a scale in which the greater magnitude has the greater measure. Such a scale has the property of direction. We thus set up a primitive or empirical scale of measurement possessing these properties. For example, the scale of temperature on some standard mercury thermometer satisfies our conditions; the bore of the tube need not be uniform, and the graduations need not be equidistant on the stem.

But when we are dealing, as we usually are, with quantities to which the notion of addition is applicable, we are soon led to abandon our empirical scale in favour of the immense advantages of a *uniform scale*, if such a scale exists. The uniform scale has the property that the measure of the sum of two members of K is the sum of the measures of the separate members. *Physical addition is represented by algebraic addition.*

It is not obvious *a priori* that a uniform scale must exist for the class K, but we shall assume that such a scale does exist for each class of quantities with which we wish to deal—for masses and lengths and intervals of time. Such quantities are called *scalar quantities*, and we are thus led to the formal definition:

The class K is a class of scalar quantities if:

(i) A member of K is completely specified by a single number (its measure) on an appropriate scale;

(ii) There is a meaning to the physical sum of two members of K, and this sum is itself a member of K;

(iii) There exists a uniform scale of measurement, i.e. a scale in which the measure of the physical sum of any two members of K is the sum of the measures of the two members.

For example, if K is the class of masses, and we use a uniform scale, a body which is formed by the cohesion of bodies of masses m_1 and m_2 (the symbols are the measures) has mass m_1+m_2. It is this property of mass, when measured on a uniform scale, that is implied in the traditional definition of mass as 'quantity of matter'. If we measure lengths on a uniform scale, and A, B, C are three points in order on a straight line, the length of the segment AC is the sum of the lengths of AB and BC. If we measure time on a uniform scale, and

A, B, C are three instants, B being later than A and C later than B, then the interval AC is the sum of the intervals AB and BC. When we come to temperature, the question is deeper; we can easily construct an empirical scale, but it is not at once obvious how to construct a uniform scale. The difficulty is that there is no immediately obvious process of addition related to the idea of temperature. But the appropriate additive process reveals itself in theoretical thermodynamics, and we construct a uniform scale—the 'absolute scale of temperature'.

1·2 Uniqueness of the uniform scale. The uniform scale, if it exists, is unique, save for a constant multiplier. To prove this, suppose that x is the measure of a member X of K on a uniform scale, and $\phi(x)$ the measure of X on a new uniform scale. (In the first instance we will suppose $x \geqslant 0$, though this restriction will later be lifted.) We may assume that $\phi(x)$ is continuous and monotonic increasing, to preserve the continuity and direction of the scale. We have

$$\phi(a+b) = \phi(a)+\phi(b),$$

for arbitrary positive (or zero) values of a and b. We first observe, by taking $a = 0$, that $\phi(0) = 0$. Next, we easily prove, by induction, that if m is a positive integer,

$$\phi(ma) = m\phi(a).$$

Thus we have $$\phi(ma-nb)+\phi(nb) = \phi(ma),$$

whence $$\phi(ma-nb) = m\phi(a)-n\phi(b),$$

where m and n are positive integers, and $ma \geqslant nb$. If we take $ma = nb$ we deduce

$$\frac{\phi(a)}{a} = \frac{\phi(b)}{b}$$

for arbitrary positive values of a and b whose ratio is rational. Thus, taking now $b = 1$, we have for positive rational values of a

$$\phi(a) = ka,$$

where k is positive since ϕ is monotonic increasing. The completion of the proof for irrational values of a follows from the continuity of ϕ.

The proof is much simpler if we make the additional assumption that ϕ is differentiable. Starting from the equation

$$\phi(a+b) = \phi(a)+\phi(b),$$

we consider variations in a and b such that $a+b$ remains constant. We deduce immediately

$$\phi'(a) = \phi'(b)$$

for arbitrary a and b. Thus

$$\phi'(a) = k,$$

whence $$\phi(a) = ka+C.$$

Finally, $k > 0$, since ϕ is increasing, and $C = 0$, since $\phi(0) = 0$.

1·3 Units. We have seen that the uniform scale is unique, save for choice of the multiplier k. We are thus led to the idea of the *unit* member of K, the member of K whose measure is 1. The unit is at our disposal, but when once it has been chosen, the scale is determined. Every member of K can be thought of as an appropriate multiple of the unit. A scalar quantity can thus be thought of as the synthesis of three elements: (i) the kind of quantity, (ii) the unit, and (iii) the measure, which is the number of units in the given quantity. For example, 'a mass of ten pounds' involves the ideas of kind (mass) and unit (pound) and number of units (ten). Strictly we ought always to distinguish between the measure (ten) and the magnitude (ten pounds) of the scalar quantity (a mass of ten pounds), though in practice a little elasticity of expression is permissible. But we must not lose sight of the fact that in the equations of mathematical physics the symbols that appear relating to scalar quantities—for example, 'a mass m', 'a time t'—denote *numbers*, the measures of the physical quantities.

If we change the unit, the measures change; if the new unit is two of the old units, the measures are reduced to one-half their former values. More generally, if the ratio of the units is $u_1 : u_2$ (i.e. the new unit has measure u_2/u_1 on the old scale) and n_1, n_2 are the measures of the same member of K on the two scales, then $n_1 u_1 = n_2 u_2$. (Cf. Chap. XXVI.)

We can construct measuring instruments with which we can determine the measures of our quantities, on a uniform scale, to a high degree of accuracy. The construction and use of such instruments is a matter that we shall not enter upon in this book.

1·4 Directed lengths. Hitherto we have thought of the measures as positive numbers, but they are not always positive. For example, electric charge may be positive or negative. The measures satisfy the condition (iii) in the definition of scalar quantities; e.g. the sum of charges $+e$ and $-e$ is a charge zero.

When we consider lengths and intervals of time we can introduce negative measures so as to include in the measures not only the idea of size but also the ideas of 'right and left' or 'before and after'. Suppose we are dealing with points on a straight line. It will be convenient to mark a fixed point O of the line as origin, and to fix on a definite sense in the line as the positive sense. If the line is drawn horizontally in the figure, we conventionally take the positive sense to the right; for the present we may treat the phrases 'B is to the right of A' and 'B is on the positive side of A' as synonymous. If X is a point on the line we label the point X with the number x, where x is the measure of the length of OX if X is to the right of (i.e. on the positive side of) O, and x is $-x'$, where x' is the measure of the length of XO, if X is to the left of O. The lengths are of course measured on a uniform scale. If A, B are two points of the line whose labels are a, b, and B is to the right of A, then the measure of the length of AB is $b-a$.

A length is strictly a positive quantity. We now introduce the idea of a *directed length*, which is not necessarily positive. If B is to the right of A the measure of the directed length of AB is simply the measure of the length of AB; if B is to the left of A the measure of the directed length of AB is minus the measure of the length of BA. We denote the measure of the directed length of AB by $\overline{AB}$. If A, B are any two points of the line, $\overline{AB} = -\overline{BA}$, and the measure of the directed length of AB is $b-a$ whether B is to the right of A or to the left. The equation

$$\overline{AB}+\overline{BC} = \overline{AC},$$

or, in symmetrical form,

$$\overline{AB}+\overline{BC}+\overline{CA}=0,$$

is valid for all positions of A, B, C on the line. Notice that our method of labelling the points of the line already involves the essential idea of a directed length. If we wish to distinguish between the directed length of AB and the ordinary length, we may denote the ordinary length by $|AB|$.

We label instants of time in the same sort of way in which we labelled points on the line. We take an instant O as origin. An instant T is labelled with the number t, where t is the measure of the interval (or, briefly, the interval) OT if T is after O, and t is $-t'$, where t' is the interval TO, if T is before O. The intervals are of course measured on a uniform scale. If A, B are two instants we define the measure of the *directed interval* AB as the measure of the interval AB if B is after A, and as minus the measure of the interval BA if B is before A. If A, B are any two instants whose labels are a, b, then $b-a$ is the measure of the directed interval AB, whether B is after A or before A. The measure $\overline{AB}$ of the directed interval AB is positive if B is after A, negative if B is before A.

The directed lengths and intervals take a place in the theory intermediate between the positive scalar quantities and the vector quantities which are the next kind of physical quantities that we consider. The essential importance of the idea of a directed length emerges when we are dealing with a set of vector quantities which are all parallel to a given line.

1·5 Directed quantities and their vectors. The next classes of physical quantities which we must consider have direction (i.e. direction in space) as well as size. Familiar examples are velocity and force. We consider a class K of such directed quantities, and in the first instance we suppose that they are not localized, i.e. size and direction together serve as a complete description of a member of K. Later on (for example, when we deal with forces acting at different points of a rigid body) we shall need to consider the extension of the theory to classes whose members are localized; in that case to describe a member of K we need to specify not only size and direction, but also the point of space with which it is associated. But for the present we deal with classes whose members are completely specified, for the purpose of the theory, by size and direction.

We assume to begin with that size and direction are independent elements in the structure of a member of K. There is a meaning to the magnitude dissociated from the direction, and we can speak of two members of K being of equal magnitude although they are not in the same direction. Thus, we can think of the members of K as quantities of the kind already discussed, which are specified on an appropriate scale by a single positive (or zero) measure, to which the idea of direction has been added. Actually, however, a slightly different analysis is more convenient. We isolate the measure of the non-directed quantity, and we call the abstract directed quantity, which is built up of this measure and the direction, the *vector* of the directed quantity. A vector is a positive (or zero) number, called the *modulus* of the vector, associated with a direction. All vectors of zero modulus are assumed to be identical; any such vector is called a null vector.

The word 'direction' as used above includes sense. In everyday speech the word 'direction' is used ambiguously, sometimes implying 'parallel to a given line', and sometimes implying 'parallel to a given line, and in a definite sense'. It is in the second of these ways that we have used the word 'direction'. This usage involves not merely the idea of a line, but of a line marked with an arrow. With this usage, if we take a sphere, centre O, every point D on the sphere defines a unique direction OD; and conversely, every direction defines a unique point on the sphere.

1·6 The vector diagram. We represent the vectors of the members of K by segments such as AB on an appropriate diagram. The diagram is to be thought of in the first instance as three-dimensional in a Euclidean space; though in most of the applications in this book we shall be concerned with classes of vectors which are all parallel to a plane. In the representation the direction AB is the direction of the vector, and the measure of the length of AB, using any convenient unit of length, is its modulus. The order in which the end-points are named is important, the direction being from A to B. Equal parallel segments in the same direction (and sense) represent the same vector, so every vector can be represented in the diagram in infinitely many ways. The representation in the diagram involves the idea of length, but this is only incidental, and only happens because length is the most convenient representation of the modulus. The notion of a vector does not essentially involve the idea of length, but only the ideas of number and direction.

We denote vectors by symbols in Clarendon type, such as $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{r}$, $\mathbf{v}$, $\mathbf{f}$; or sometimes by symbols such as $\overline{AB}$, where AB is the segment representing the vector in the diagram. The Clarendon type may be regarded as the standard notation, but the other notation is also useful, especially in the early stages, when we wish to refer frequently to the diagram. We denote the modulus of $\mathbf{P}$ by $|\mathbf{P}|$. The nul vector is denoted by $\mathbf{0}$.

1·7 Vector quantities. The most important kind of directed quantities are those called *vector quantities*. Vector quantities satisfy a condition which is related to the process of addition, and is analogous to the requirement of a uniform scale for scalar quantities. We consider now the formal definition.

The class K is a class of vector quantities if:

(i) A member of K is completely specified by its vector;

(ii) There is a meaning to the physical sum of two members of K, and this sum is itself a member of K;

(iii) There exists a scale of measurement such that, if the vectors of two members of K are $\overline{AB}$ and $\overline{BC}$, the vector of their sum is $\overline{AC}$.

The process described in (iii) is called *vector summation*, and the vector formed from two given vectors $\mathbf{P}$ and $\mathbf{Q}$ by this process (i.e. the vector $\overline{AC}$ if $\overline{AB}$ represents $\mathbf{P}$ and $\overline{BC}$ represents $\mathbf{Q}$) is called their vector sum. It is denoted by $\mathbf{P}+\mathbf{Q}$. We can express the definition informally by saying that directed quantities are vector quantities if physical summation is represented by vector summation.

Notice that we speak of the 'vector' of a directed quantity even when this quantity is not a vector quantity. Actually the directed quantities we deal with in this book will all be vector quantities, so the slightly awkward nomenclature will not give rise to any inconvenience.

1·8 Parallel vectors. Some implications of the definition of a vector quantity should be noticed. Suppose first that we have a set of vectors all parallel and all in the same sense. Then the modulus of the sum of two vectors is the sum of their moduli. The vectors reduce effectively to numbers, the vector quantities to scalar quantities. Physical summation is represented by algebraic summation, and we see again that the uniform scale of measurement is fundamental.

Next, consider a set of vectors all parallel to a line λ, but not all in the same sense. Now the given line is associated with *two* directions, if we use the word 'direction' to include sense, as we have done hitherto. But it is usually convenient to fix on one definite sense in λ as the positive sense, and to describe a vector $\overline{AB}$ (where A and B are two points of λ) by the directed length of AB. Indeed, we have already anticipated this representation by using the same notation $\overline{AB}$ both for the vector and for the directed length. If A, B, C are any three points of λ the equation

$$\overline{AB}+\overline{BC} = \overline{AC}$$

is valid whether we interpret a symbol such as $\overline{AB}$ as a vector or as a directed length. With this representation, physical addition corresponds to algebraic addition.

The modulus with the appropriate sign is called the *directed modulus*. Thus the directed modulus is equal to the modulus if the vector is in the positive sense of λ, and to the modulus with the minus sign prefixed if the vector is in the negative sense of λ. A set of vector quantities all parallel to a line λ, and represented by their directed moduli, can be treated in practice as a set of scalar quantities.

When we use directed moduli we sacrifice the advantage of uniqueness in the description of the vectors. Every vector has a definite (positive) modulus, and a definite direction (including sense). But when we use directed moduli we can describe the same vector in two ways, since a vector $\mathbf{P}$ in the negative sense of λ is also described as a vector $-\mathbf{P}$ in the positive sense.

However, no confusion will arise from this, and the notation has great advantages when we deal with a set of vectors all parallel to the same line. Perhaps the advantages are most conspicuous when we use rectangular axes, and we find it convenient to fix on a definite sense in each axis as the positive sense.

1·9 The commutative property of vector summation. It is clear from elementary geometry, since equal parallel segments in the diagram represent the same vector, that vector summation is commutative,

$$\mathbf{P}+\mathbf{Q} = \mathbf{Q}+\mathbf{P}.$$

This equation is illustrated in Fig. 1·9.

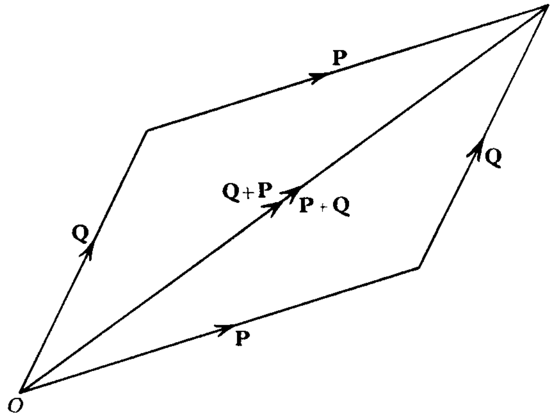

Fig. 1·9

Since $$\overline{AB}+\overline{BA} = \mathbf{0},$$

it is natural to define the negative of a given vector by the equation

$$\overline{AB} = -\overline{BA},$$

and to define subtraction by the equation

$$\mathbf{P}-\mathbf{Q} = \mathbf{P}+(-\mathbf{Q}).$$

We can preserve the analogy with directed lengths, even when the vectors are not all parallel, by writing the rule for vector summation in the symmetrical form

$$\overline{AB}+\overline{BC}+\overline{CA} = \mathbf{0}.$$